Escola de
ECONOMIA
de São Paulo

Textos para Discussão

250

Março de 2010

A Simple Demonstration of the existence of the Jordan Canonical Form for any square matrix

Robert Norman V. Cajado Nicol

FGV
EESP

A Simple Demonstration of the existence of the Jordan Canonical Form for any square matrix

Robert N Cajado Nicol
FGV/EESP 22 nd March 2010

Abstract

All the demonstrations known to this author of the existence of the Jordan Canonical Form are somewhat complex - usually invoking the use of new spaces, and what not. These demonstrations are usually too difficult for an average Mathematics student to understand how he or she can obtain the Jordan Canonical Form for any square matrix. The method here proposed not only demonstrates the existence of such forms but, additionally, shows how to find them in a step by step manner. I do not claim that the following demonstration is in any way "elegant" (by the standards of elegance in fashion nowadays among mathematicians) but merely simple (undergraduate students taking a fist course in Matrix Algebra would understand how it works).

Sumário

Todas as demonstrações conhecidas por este autor da existência da Forma Canônica de Jordan são por demais complexas – envolvendo o uso de novos espaços e outras pirotecnias. Tais demonstrações são normalmente difíceis para um aluno médio de Matemática entender. O método aqui proposto não somente demonstra a existências dessas formas, como mostra uma forma simples com poucos passos para obtê-las. A demonstração não pretende ser "elegante", mas tão somente, simples.

JEL : C02, C65.

First Step

We start with a known result: any square matrix can be transformed into an upper (or lower) triangular matrix by means of a similarity transformation.

Suppose we have an $\boldsymbol{n}_{\times}\boldsymbol{n}$ matrix $\boldsymbol{A}$

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

we can always find the neigenvalues corresponding to the solution to the determinantal equation

$$Det(A - \lambda I) = Det\begin{pmatrix} (a_{11} - \lambda) & \cdots & (a_{1n} - \lambda) \\ \vdots & \ddots & \vdots \\ (a_{n1} - \lambda) & \cdots & (a_{nn} - \lambda) \end{pmatrix} = 0$$

Suppose these are:

$$\lambda_1, \lambda_2, \ldots \lambda_n$$

Corresponding to each one of these $\boldsymbol{n}$ eigenvalues we can find $\boldsymbol{n}$ eigenvectors

$$x_1, x_2, \ldots x_n$$

Let us start with the first root λ_1 and suppose that the corresponding right side eigenvector is x_1'. We can always normalize this vector so as to have $x_1 x_1' = 1$. So far we have got to the stage where we have

$$Ax_1' = \lambda_1 x_1' \quad \text{or} \quad A x_1' = x_1' \lambda_1 \qquad \text{with} \quad x_1 x_1' = 1$$

Starting with x_1' we can always build an orthonormal basis for our space R^n. Let this set be $x_1', w_2', w_3', \ldots w_n'$. Let

$$W = \begin{pmatrix} x_1 \\ w_2 \\ w_3 \\ \vdots \\ w_n \end{pmatrix} \quad and \quad W' = \begin{pmatrix} x_1' & w_2' & w_3' & w_4' & \cdots & w_n' \end{pmatrix}$$

Then

$$WAW' = \begin{pmatrix} x_1 A x_1' & x_1 A w_2' & \cdots & x_1 A w_n' \\ w_2 A x_1' & w_2 A w_2' & \cdots & w_2 A w_n' \\ \vdots & \vdots & \ddots & \vdots \\ w_n A x_1' & w_n A w_2' & \cdots & w_n A w_n' \end{pmatrix} = \begin{pmatrix} x_1 x_1' \lambda_1 & x_1 A w_2' & \cdots & x_1 A w_n' \\ w_2 x_1' \lambda_1 & w_2 A w_2' & \ldots & w_2 A w_n' \\ \vdots & \vdots & \ddots & \vdots \\ w_n x_1' \lambda_1 & w_n A w_2' & \cdots & w_n A w_n' \end{pmatrix}$$

But we know that $\quad x_1 x_1' = 1 \quad w_2 x_1' = 0 \cdots \quad w_n x_1' = 0 \qquad$ hence

$$WAW' = \begin{pmatrix} \lambda_1 & x_1 A w_2' & \cdots & x_1 A w_n' \\ 0 & w_2 A w_2' & \cdots & w_2 A w_n^1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & w_n A w_2' & \cdots & w_n A w_n^1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

And now we can restart the process by finding out the (n-1) eigenvalues and the corresponding right eigenvectors for the matrix

$$B = \begin{pmatrix} b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

Suppose one of these eigenvectors is λ_2 and that the corresponding right side eigenvector is s_2' (this eigenvector being of size (n-1)x1), that is

$$Bs_2' = \lambda_2 s_2'$$

And starting with this eigenvector we can build an orthonormal basis for the space R^{n-1}.
Let such a basis be s_2' , u_3', u_4', … u_n' . Then by pre-multiplying and post-multiplying the matrix **WAW'** by **U** and **U'** where

$$U=\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & & s_2 & & \\ 0 & & u_3 & & \\ 0 & & u_4 & & \\ 0 & & u_5 & & \\ & & \ddots & & \\ 0 & & u_n & & \end{pmatrix}$$

We will get

$$UWAW'U'=\begin{pmatrix} \lambda_1 & b_{12} & b_{13} & \cdots & b_{1n} \\ 0 & \lambda_2 & c_{33} & \cdots & c_{1n} \\ 0 & 0 & c_{43} & \cdots & c_{4n} \\ 0 & 0 & c_{53} & \cdots & c_{5n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & c_{n3} & \cdots & c_{nn} \end{pmatrix}$$

Repeating the process n-2 times we should get an upper triangular matrix with all the eigenvectors of A in the main diagonal. Furthermore if we take care of selecting the eigenvalues that will give us the eigenvectors for the similarity transformation that we will end up with in such a way as to get all the eigenvectors connected with the roots that have multiplicity higher than one in a sequence we should end up with un upper triangular matrix of the form

$$F=PAP'=\begin{pmatrix} \lambda & f_{12} & f_{13} & f_{14} & f_{15} & \cdots & f_{1(n-1)} & f_{1n} \\ 0 & \lambda & f_{23} & f_{24} & f_{25} & \cdots & f_{2(n-1)} & f_{2(n-1)} \\ 0 & 0 & \theta & f_{34} & f_{35} & \cdots & f_{3(n-1)} & f_{3n} \\ 0 & 0 & 0 & \theta & f_{45} & \cdots & f_{4(n-1)} & f_{4n} \\ 0 & 0 & 0 & 0 & \theta & \cdots & f_{5(n-1)} & f_{5n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & f_{(n-1)n} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi \end{pmatrix} \quad where \quad PAP'=I$$

All this is well known.

Second Step

If we now use the similarity transformation ***(I-k) F (I+k)*** where

$$(I-k)=\begin{pmatrix} 1 & 0 & -k & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad and \quad (I+k)=\begin{pmatrix} 1 & 0 & k & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

We can see: first that ***(I-k)(I+k) = I***

And, secondly that

$$(I-k)A(I+k)=\begin{pmatrix} \lambda & f_{12} & (f_{13}-k\theta) & (f_{14}-kf_{34}) & (f_{15}-kf_{35}) & \cdots & (f_{1(n-1)}-kf_{3(n-1)}) & (f_{1n}-kf_{3n}) \\ 0 & \lambda & f_{23} & f_{24} & f_{25} & \cdots & f_{2(n-1)} & f_{2(n-1)} \\ 0 & 0 & \theta & f_{34} & f_{35} & \cdots & f_{3(n-1)} & f_{3n} \\ 0 & 0 & 0 & \theta & f_{45} & \cdots & f_{4(n-1)} & f_{4n} \\ 0 & 0 & 0 & 0 & \theta & \cdots & f_{5(n-1)} & f_{5n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & f_{(n-1)n} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi \end{pmatrix}(I+k)$$

$$=\begin{pmatrix} \lambda & f_{12} & (f_{13}-k\theta+k\lambda) & (f_{14}-kf_{34}) & (f_{15}-kf_{35}) & \cdots & (f_{1(n-1)}-kf_{3(n-1)}) & (f_{1n}-kf_{3n}) \\ 0 & \lambda & f_{23} & f_{24} & f_{25} & \cdots & f_{2(n-1)} & f_{2(n-1)} \\ 0 & 0 & \theta & f_{34} & f_{35} & \cdots & f_{3(n-1)} & f_{3n} \\ 0 & 0 & 0 & \theta & f_{45} & \cdots & f_{4(n-1)} & f_{4n} \\ 0 & 0 & 0 & 0 & \theta & \cdots & f_{5(n-1)} & f_{5n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & f_{(n-1)n} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi \end{pmatrix}$$

And the third term of the first row that is $(f_{13}-k\theta+k\lambda)$ can be made equal to zero provided we choose an adequate value for ***k*** . We can see that $k=\frac{f_{13}}{\theta-\lambda}$ would do the trick. But that would only work iff $\theta \neq \lambda$.

So what we end up with by using this type of transformation for getting rid of all possible terms off the main diagonal would be a matrix of the type:

$$M=\begin{pmatrix} \lambda & f_{12} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \theta & f_{34} & f_{35} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \theta & f_{45} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \phi & f_{67} & \cdots & f_{6(n-1)} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & \cdots & f_{7(n-1)} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \phi & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \sigma \end{pmatrix} \quad \text{or}$$

$$M = \begin{pmatrix} B & & & \\ & C & & \\ & & D & \\ & & & E \end{pmatrix} \quad \text{where} \quad B = \begin{pmatrix} \lambda & f_{12} \\ 0 & \lambda \end{pmatrix} \quad C = \begin{pmatrix} \theta & f_{34} & f_{35} \\ 0 & \theta & f_{45} \\ 0 & 0 & \theta \end{pmatrix}$$

$$D = \begin{pmatrix} \phi & f_{67} & \cdots & f_{6(n-1)} \\ 0 & \phi & \cdots & f_{7(n-1)} \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \phi \end{pmatrix} \quad \text{and} \quad E = (\delta)$$

Where ***M*** is an upper triangular matrix composed of several smaller upper triangular matrices ***B, C, D*** and ***E***.
These smaller matrices ***B, C*** and ***D*** have in their main diagonal only identical terms that correspond in value to the roots with multiplicity greater than one such as λ, θ, ϕ and (in the example above) and one matrix $\boldsymbol{E_{1x1}}$ which has only one term ***(σ)*** that correspond to the only root that has multiplicity equal to one (in this example).

So the problem of finding a similarity transformation that will transform a matrix ***A*** into a Jordan Canonical Form, boils down to finding similarity transformations that will transform upper triangular matrices similar to the matrices ***B, C , D*** and ***E*** into their corresponding ***JCFs***.

Third Step

Transforming a matrix of the type

$$D = \begin{pmatrix} \phi & f_{67} & \cdots & f_{6(n-1)} \\ 0 & \phi & \cdots & f_{7(n-1)} \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \phi \end{pmatrix} \text{ into its } \boldsymbol{JCF}$$

For the sake of clarity let me suppose that ***D*** is a ***6***x***6*** matrix, that is that

$$D = \begin{pmatrix} \phi & a & b & c & d & e \\ 0 & \phi & f & g & h & i \\ 0 & 0 & \phi & j & k & l \\ 0 & 0 & 0 & \phi & m & n \\ 0 & 0 & 0 & 0 & \phi & p \\ 0 & 0 & 0 & 0 & 0 & \phi \end{pmatrix}$$

Let us assume for the moment that the value of

$$\phi = 1$$

We will see that this assumption can be dropped in the end and we will get the

$$\phi \neq 1$$

back.

So, for the moment we are assuming that

$$D = \begin{pmatrix} 1 & a & b & c & d & e \\ 0 & 1 & f & g & h & i \\ 0 & 0 & 1 & j & k & l \\ 0 & 0 & 0 & 1 & m & n \\ 0 & 0 & 0 & 0 & 1 & p \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let us start by transforming the superdiagonal (upper diagonal neighbor of main diagonal) into a series of ones (that is the diagonal with the terms ***a, f, j, m , p***) This is done quite simply by post-multiplying ***D*** by the matrix by the matrix ***S*** and pre-multiplying it by its inverse .

$$DS = \begin{pmatrix} 1 & a & b & c & d & e \\ 0 & 1 & f & g & h & i \\ 0 & 0 & 1 & j & k & l \\ 0 & 0 & 0 & 1 & m & n \\ 0 & 0 & 0 & 0 & 1 & p \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{p} \end{pmatrix} = \begin{pmatrix} 1 & a & b & c & d & \frac{e}{p} \\ 0 & 1 & f & g & h & \frac{i}{p} \\ 0 & 0 & 1 & j & k & \frac{l}{p} \\ 0 & 0 & 0 & 1 & m & \frac{n}{p} \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{p} \end{pmatrix}$$

and

$$S^{-1}DS = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & p \end{pmatrix} \begin{pmatrix} 1 & a & b & c & d & \frac{e}{p} \\ 0 & 1 & f & g & h & \frac{i}{p} \\ 0 & 0 & 1 & j & k & \frac{l}{p} \\ 0 & 0 & 0 & 1 & m & \frac{n}{p} \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{p} \end{pmatrix} = \begin{pmatrix} 1 & a & b & c & d & \frac{e}{p} \\ 0 & 1 & f & g & h & \frac{i}{p} \\ 0 & 0 & 1 & j & k & \frac{l}{p} \\ 0 & 0 & 0 & 1 & m & \frac{n}{p} \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let us now proceed with the m term in a similar fashion post-multiplying by ***T*** and pre-multiplying by its inverse, where the meaning of ***T*** is easy to see.

$$S^{-1}DST = \begin{pmatrix} 1 & a & b & c & d & \frac{e}{p} \\ 0 & 1 & f & g & h & \frac{i}{p} \\ 0 & 0 & 1 & j & k & \frac{l}{p} \\ 0 & 0 & 0 & 1 & m & \frac{n}{p} \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{m} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{m} \end{pmatrix} = \begin{pmatrix} 1 & a & b & c & \frac{d}{m} & \frac{e}{mp} \\ 0 & 1 & f & g & \frac{h}{m} & \frac{i}{mp} \\ 0 & 0 & 1 & j & \frac{k}{m} & \frac{l}{mp} \\ 0 & 0 & 0 & 1 & 1 & \frac{n}{mp} \\ 0 & 0 & 0 & 0 & \frac{1}{m} & \frac{1}{m} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{m} \end{pmatrix}$$

$$T^{-1}S^{-1}DST = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & m & 0 \\ 0 & 0 & 0 & 0 & 0 & m \end{pmatrix} \begin{pmatrix} 1 & a & b & c & \frac{d}{m} & \frac{e}{mp} \\ 0 & 1 & f & g & \frac{h}{m} & \frac{i}{mp} \\ 0 & 0 & 1 & j & \frac{k}{m} & \frac{l}{mp} \\ 0 & 0 & 0 & 1 & 1 & \frac{n}{mp} \\ 0 & 0 & 0 & 0 & \frac{1}{m} & \frac{1}{m} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{m} \end{pmatrix} = \begin{pmatrix} 1 & a & b & c & \frac{d}{m} & \frac{e}{mp} \\ 0 & 1 & f & g & \frac{h}{m} & \frac{i}{mp} \\ 0 & 0 & 1 & j & \frac{k}{m} & \frac{l}{mp} \\ 0 & 0 & 0 & 1 & 1 & \frac{n}{mp} \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let us now proceed to j. It can be seen quite easily that by using U and its inverse

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{j} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{j} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{j} \end{pmatrix} \qquad U^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & j & 0 & 0 \\ 0 & 0 & 0 & 0 & j & 0 \\ 0 & 0 & 0 & 0 & 0 & j \end{pmatrix}$$

We should get a matrix of the form

$$\begin{pmatrix} 1 & a & b & \frac{c}{j} & \frac{d}{jm} & \frac{e}{jmp} \\ 0 & 1 & f & \frac{g}{j} & \frac{h}{jm} & \frac{i}{jmp} \\ 0 & 0 & 1 & 1 & \frac{k}{jm} & \frac{l}{jmp} \\ 0 & 0 & 0 & 1 & 1 & \frac{n}{jmp} \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

After another step to deal with f and a we get

$$\begin{pmatrix} 1 & a & \frac{b}{f} & \frac{c}{fj} & \frac{d}{fjm} & \frac{e}{fjmp} \\ 0 & 1 & 1 & \frac{g}{fj} & \frac{h}{fjm} & \frac{i}{fjmp} \\ 0 & 0 & 1 & 1 & \frac{k}{fjm} & \frac{l}{fjmp} \\ 0 & 0 & 0 & 1 & 1 & \frac{n}{fjmp} \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \textit{and} \quad \textit{finally} \quad \begin{pmatrix} 1 & 1 & \frac{b}{af} & \frac{c}{afj} & \frac{d}{afjm} & \frac{e}{afjmp} \\ 0 & 1 & 1 & \frac{g}{afj} & \frac{h}{afjm} & \frac{i}{afjmp} \\ 0 & 0 & 1 & 1 & \frac{k}{afjm} & \frac{l}{afjmp} \\ 0 & 0 & 0 & 1 & 1 & \frac{n}{afjmp} \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

One can see that the whole process could have been performed in one single step by post-multiplying and pre-multiplying matrix ***D*** by ***P*** and its inverse

$$P=\begin{pmatrix}1&0&0&0&0&0\\0&\frac{1}{a}&0&0&0&0\\0&0&\frac{1}{af}&0&0&0\\0&0&0&\frac{1}{afj}&0&0\\0&0&0&0&\frac{1}{afjm}&0\\0&0&0&0&0&\frac{1}{afjmp}\end{pmatrix}\quad and\quad P^{-1}=\begin{pmatrix}1&0&0&0&0&0\\0&a&0&0&0&0\\0&0&af&0&0&0\\0&0&0&afj&0&0\\0&0&0&0&afjm&0\\0&0&0&0&0&afjmp\end{pmatrix}$$

Let us call the end product of our transformation matrix ***J*** and let us simplify our notation

$$J=\begin{pmatrix}1&1&\alpha&\beta&\chi&\delta\\0&1&1&\varepsilon&\phi&\varphi\\0&0&1&1&\gamma&\eta\\0&0&0&1&1&\kappa\\0&0&0&0&1&1\\0&0&0&0&0&1\end{pmatrix}$$

Let us recall that the only assumption that we have made so far is that the diagonal element in the original matrix ***D***, that is elements such as ***a, f, j, m, p*** were all different from zero. We will see later that this is an assumption that is not necessary.
Let us star by eliminating $\boldsymbol{\alpha}$ from matrix ***J*** . This can be done as follows

$$J\mathrm{A}=\begin{pmatrix}1&1&\alpha&\beta&\chi&\delta\\0&1&1&\varepsilon&\phi&\varphi\\0&0&1&1&\gamma&\eta\\0&0&0&1&1&\kappa\\0&0&0&0&1&1\\0&0&0&0&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0&0&0\\0&1&-\alpha&0&0&0\\0&0&1&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{pmatrix}=\begin{pmatrix}1&1&0&\beta&\chi&\delta\\0&1&(1-\alpha)&\varepsilon&\phi&\varphi\\0&0&1&1&\gamma&\eta\\0&0&0&1&1&\kappa\\0&0&0&0&1&1\\0&0&0&0&0&1\end{pmatrix}$$

$$A^{-1}JA=\begin{pmatrix}1&0&0&0&0&0\\0&1&\alpha&0&0&0\\0&0&1&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{pmatrix}\begin{pmatrix}1&1&0&\beta&\chi&\delta\\0&1&(1-\alpha)&\varepsilon&\phi&\varphi\\0&0&1&1&\gamma&\eta\\0&0&0&1&1&\kappa\\0&0&0&0&1&1\\0&0&0&0&0&1\end{pmatrix}=\begin{pmatrix}1&1&0&\beta&\chi&\delta\\0&1&1&(\varepsilon+\alpha)&(\phi+\alpha\gamma)&(\varphi+\alpha\eta)\\0&0&1&1&\gamma&\eta\\0&0&0&1&1&\kappa\\0&0&0&0&1&1\\0&0&0&0&0&1\end{pmatrix}$$

Let us call this new matrix ***J**** and simplify our notation

$$J^* = \begin{pmatrix} 1 & 1 & 0 & \beta & \chi & \delta \\ 0 & 1 & 1 & \varepsilon' & \phi' & \varphi' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let us now eliminate $\boldsymbol{\beta}$

$$J^*\mathrm{B} = \begin{pmatrix} 1 & 1 & 0 & \beta & \chi & \delta \\ 0 & 1 & 1 & \varepsilon' & \phi' & \varphi' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -\beta & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & \chi & \delta \\ 0 & 1 & 1 & (\varepsilon'-\beta) & \phi' & \varphi' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathrm{B}^{-1}J^*\mathrm{B} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \beta & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 & \chi & \delta \\ 0 & 1 & 1 & (\varepsilon'-\beta) & \phi' & \varphi' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & \chi & \delta \\ 0 & 1 & 1 & \varepsilon' & (\phi'+\beta) & (\varphi'+\beta\kappa) \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let us now eliminate χ from the new matrix ***J*****

$$J^{**}\mathrm{X} = \begin{pmatrix} 1 & 1 & 0 & 0 & \chi & \delta \\ 0 & 1 & 1 & \varepsilon' & \phi'' & \varphi'' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -\chi & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & \delta \\ 0 & 1 & 1 & \varepsilon' & (\phi''-\chi) & \varphi'' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathrm{X}^{-1}J^{**}\mathrm{X} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \chi & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & \delta \\ 0 & 1 & 1 & \varepsilon' & (\phi''-\chi) & \varphi'' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & \delta \\ 0 & 1 & 1 & \varepsilon' & \phi'' & (\varphi''+\chi) \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The last element from the first row that we want to eliminate is δ something we can achieve by post-multiplying ***J ****** by $\boldsymbol{\Delta}$ and pre-multiplying by its inverse

$$J^{***}\Delta = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & \delta \\ 0 & 1 & 1 & \varepsilon' & \phi'' & \varphi''' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -\delta \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & \varepsilon' & \phi'' & (\varphi'''-\delta) \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Delta^{-1}J^{***}\Delta == \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \delta \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & \varepsilon' & \phi'' & (\varphi'''-\delta) \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & \varepsilon' & \phi'' & \varphi''' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let us call the new matrix instead of ***J******* something more simple such as ***K*** .
It can be seen that all the elements of the first row that we wanted to reduce to zero were in fact reduced to zero by the following similarity transformation

$$S^{-1}JS == \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \alpha & \beta & \chi & \delta \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \alpha & \beta & \chi & \delta \\ 0 & 1 & 1 & \varepsilon & \phi & \varphi \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\alpha & -\beta & -\chi & -\delta \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = K$$

where

$$K = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & \varepsilon' & \phi'' & \varphi''' \\ 0 & 0 & 1 & 1 & \gamma & \eta \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let us therefore apply the similarity transformation ***T*** and its inverse to ***K***, where

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\varepsilon' & -\phi'' & -\varphi''' \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad and \quad T^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \varepsilon' & \phi'' & \varphi''' \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

which should give us a new matrix ***K’*** such that

$$K' = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & \gamma' & \eta'' \\ 0 & 0 & 0 & 1 & 1 & \kappa \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

And proceeding in a similar way for the remaining terms off the main diagonal and the superdiagonal we will finally get what we wanted in the first place that is

$$W^{-1}DW = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Important Observations:

We assumed that the main diagonal and the terms in the superdiagonal were all different from zero. Supposing they were not? Supposing we had something like the matrix ***A (4x4)*** below

$$A = \begin{pmatrix} 0 & 0 & a & b \\ 0 & 0 & c & d \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

What should we do? Simply add ***(I + H)*** to get $\hat{A}$ find the ***JCF*** and subtract from the final result ***(I+H)***
where

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad and \quad H = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The similarity transformation will not affect the sum of these two matrices so that in the end of the process we will still have these two matrices to subtract from our result.
What I am saying is that supposing the similarity transformation was represented by the matrix P and its inverse, then

$$P^{-1}AP = P^{-1}(A + [I + H] - [I + H])P = P^{-1}\hat{A}P - P^{-1}[I + H]P = P^{-1}\hat{A}P - [I + H]$$

A second observation related to this last one. We supposed that all the roots of our matrix ***D*** had values equal to 1. Well supposing that their real value was $\lambda \neq 1$. How do we deal with that? We would use a trick similar to this last one, that is we would add and subtract a diagonal matrix ***F*** such that

$$F = \begin{pmatrix} (1-\lambda) & 0 & 0 & 0 & 0 & 0 \\ 0 & (1-\lambda) & 0 & 0 & 0 & 0 \\ 0 & 0 & (1-\lambda) & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-\lambda) & 0 & 0 \\ 0 & 0 & 0 & 0 & (1-\lambda) & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-\lambda) \end{pmatrix}$$

and we would see that the similarity transformation applied to this new matrix ***D'*** would leave ***F*** unaltered so that in the end the same ***F*** that was initially added could be subtracted.

Having worked out the similarity transformation for all the submatrices of the original matrix ***M*** where

$$M = \begin{pmatrix} B & & & \\ & C & & \\ & & D & \\ & & & E \end{pmatrix}$$

we would have something like

$$\bar{P}^{-1}M\bar{P} = \begin{pmatrix} P_B^{-1} & & & \\ & P_C^{-1} & & \\ & & P_D^{-1} & \\ & & & P_E^{-1} \end{pmatrix}\begin{pmatrix} B & & & \\ & C & & \\ & & D & \\ & & & E \end{pmatrix}\begin{pmatrix} P_B & & & \\ & P_C & & \\ & & P_D & \\ & & & P_E \end{pmatrix} = \begin{pmatrix} P_B^{-1}BP_B & & & \\ & P_C^{-1}CP_C & & \\ & & P_D^{-1}DP_D & \\ & & & P_E^{-1}EP_E \end{pmatrix}$$

Which would be the JCF we were looking for.
I believe I have delivered what I had promised at the very start of these 12 pages.

Comments are welcome.